A CRITIQUE OF THE WTO JURISPRUDENCE ON 'NECESSITY'

GISELE KAPTERIAN*

Abstract This paper examines whether the evolving jurisprudence on necessity as developed by the WTO adjudicatory bodies reflects the same balance between trade liberalization and regulatory autonomy as that contained in the WTO treaty texts, particularly with regard to the GATT. It is argued that a divergence can be observed which raises questions of competence, legitimacy and transparency. Specific amendments to the prevailing test are also proposed in order to achieve what the author suggests is a textually consistent, and thus legitimate, necessity test of equal efficacy.

I. INTRODUCTION

A significant and ongoing challenge facing the WTO system is how to balance the pursuit of free trade with the need to afford sufficient space for the domestic regulatory autonomy of its Members. The various actions taken by Members choosing to promote national policy objectives purportedly aimed at addressing non-trade issues such as environmental protection, labour rights, human health and consumer protection (often in response to the demands of the domestic constituency) have reignited the ongoing battle. The WTO Agreement simultaneously aims to encourage trade liberalization by condemning measures that illegitimately advantage domestic industry, while specifically carving out multiple exceptions to allow sufficient space for the pursuit of other important policy objectives. The factual basis for this difficulty lies in the fact that a Member's decision to regulate often advantages domestic industry, though the measure may not necessarily have been drafted with any protectionist intent. Thus, the issue facing the WTO adjudicatory bodies is whether the measure is legitimately aimed at non-trade goals with unintentional but unavoidable consequences for international trade or a disguised attempt to protect or boost domestic industry. The issue has perhaps gained more importance recently in light of the current economic crisis, to

* Associate, White & Case LLP, Geneva, Switzerland. This article was originally written and submitted as the thesis requirement in the Law of the WTO paper taken as part of the LL.M. programme at the University of Cambridge, United Kingdom. The author wishes to thank Margaret Young for her invaluable insights when supervising this paper, Prof Peter Gillies and David Hartridge for their excellent comments on an earlier draft, and Tatiana Falcao for her infinitely helpful review.

[*ICLQ* vol 59, January 2010 pp 89–127] doi:10.1017/S0020589309990091

which many Members have responded by enacting measures aimed at supporting domestic industry crucial to the survival of their domestic economies.

It is clear that the primary tool selected by the Members to distinguish between illegitimate protectionist measures and legitimate exceptions, and thus balance the dual objectives of increased trade liberalization and allowing Members sufficient space to pursue non-trade objectives, is the necessity test. Appearing in slightly different forms throughout the various agreements,[1] necessity tests have been used to reflect the extent of the Members' political compromise depending on the object and purpose of the agreement, expanding the scope afforded to domestic autonomy in some, while restricting it in others. However, the application of these tests by the WTO adjudicatory bodies has raised concern over the degree of deference afforded to Members' legitimate policy objectives, raising questions of legitimacy, competence and transparency.[2]

I will explore the question of whether this evolving jurisprudence reflects the same balance between trade liberalization and regulatory autonomy as that contained in the treaty texts, particularly with regard to the GATT. I will focus on the extent to which the adjudicating bodies' application of certain doctrinal tools when interpreting necessity has influenced the freedom afforded to domestic regulatory choices when faced with their negative trade consequences. Attention will also be paid to the parallel development emerging from the application of these doctrinal tools providing the adjudicating bodies with broader discretionary power upon which the survival of the domestic policy choice will depend. As such, the examination will reveal the extent to which the horizontal, ideological struggle between trade and non-trade values is related to the vertical, institutional struggle between the adjudicating bodies and Members.

Though my primary focus will be the interpretation of necessity under the GATT, the analysis is equally applicable to the other agreements for three reasons: first, despite the different textual provisions of necessity within and amongst the agreements, the adjudicating bodies' interpretation reveals heavy

[1] The tests that have received the most attention thus far and will be considered in this discussion are: General Agreement on Tariffs and Trade (GATT 1947), in *WTO*, *The Legal Texts: The Results of the Uruguay Round of Multilateral Trade Negotiations* (2007) 423–493; (*Legal Texts*) arts XX and XXI; Agreement on the Application of Sanitary and Phytosanitary Measures (15 April 1994); Marrakesh Agreement Establishing the WTO (hereinafter WTO Agreement), Annex 1A, Legal Texts, 59–73, arts 2.2 and 5.6 (hereinafter SPS Agreement); Agreement on Technical Barriers to Trade (15 April 1994) WTO Agreement, Annex 1A, Legal Texts 121–143, arts 2.2 and 2.5 (hereinafter TBT Agreement); General Agreement on Trade in Services, WTO Agreement, Annex 1B, Legal Texts, 284–320, arts XIV and VI:4 (hereinafter GATS Agreement).

[2] See for example, S Charnovitz, 'Environment and Health Under WTO Dispute Settlement' (1998) 32 Int'l Lawyer 901, 920–21 (highlighting environmentalists' distrust of the WTO dispute settlement system); R Howse and E Turk, 'The WTO Impact on Internal Regulations—A Case Study of the Canada-EC Asbestos Dispute,' in G de Burca and J Scott (eds), *The EU and the WTO: Legal and Constitutional Issues* (Hart Publishing, Oxford, 2003) 283; C Button, *The Power to Protect: Trade, Health and Uncertainty in the WTO* (Hart Publishing, Oxford, 2004).

cross-fertilization, making the examination of one test in isolation nonsensical if not impossible; secondly, ascertaining what was intended to be the relevant balance under one agreement is assisted by reference to the balance contained in the other agreements by way of comparison; finally, the different balances struck under each agreement are crucial to determining the broader equilibrium embodied by the WTO agreement as a whole.

I will suggest that the meaning of necessity as interpreted by the adjudicating bodies has, until recently, demonstrated increasing divergence from the language of the treaty text, and needlessly curtailed the domestic regulatory freedom afforded to Members under the treaty. Beginning with the importation of the least restrictive means (here in after LRM) test in *US-Section 337*, the GATT and WTO adjudicatory bodies have applied the necessity test to strike out the legitimate use of a domestic measure if any less trade-restrictive measure is deemed to be available, regardless of whether the alternative measure is economically feasible to implement or whether it achieves the level of protection chosen by the Member. The creation of the balancing test, first enunciated in *Korea-Beef*[3] and now firmly entrenched as the proper methodology an adjudicating body should employ when applying the necessity test, further tipped the balance in favour of trade liberalization. As subsequent applications of this test reveal, the balancing test expands the jurisdiction of the adjudicating bodies, demonstrating a disconcerting dependence on their discretion for the survival of domestic regulatory choices. The jurisprudence reveals a strong tendency to judge the value of the policy goal using the adjudicating bodies' own value system and opaque reasoning on how the elements of the balancing test interact when applied to the particular circumstances of the case. Substantial cross-fertilization of the necessity tests appearing in different agreements has further promoted the creation of a GATT necessity test at odds with the language of the text.

My primary assertion is that the balancing test does not reflect the balance of rights and obligations reached by the Members. I suggest that the aim of exposing illegitimate protectionist measures can and should be achieved through a more textually consistent interpretation of necessity. Giving proper effect to the content of article XX sub-clauses and chapeau, as well as the differences between the necessity provisions in different agreements, will provide the necessary scrutiny of Members' regulations without illegitimate intervention into regulatory autonomy. Such a test will preserve the negotiated balance between autonomy and free trade and provide a more transparent and certain legal standard against which measures should be assessed. Part II below will address the negotiated balance agreed upon by the Members with particular regard to the nature of the WTO agreement and will set out the doctrinal tools often employed in the

[3] *Korea-Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161.169/AB/R, 11 December 2000.

interpretation of necessity. Part III will describe and critique the interpretative evolution of necessity under each agreement by the adjudicating bodies, focusing on the significant developments and degrees of convergence and difference. The GATT article XX test will be compared to the emerging jurisprudence on necessity under the GATS, SPS and TBT.[4] Part IV will examine the three limbs of the *Korea-Beef* test in light of their interpretative legitimacy under the Vienna Convention on the Law of Treaties[5] (VCLT), propose and examine an alternative test and briefly examine whether LRM and balancing tests are compatible with the purpose and current structure of the WTO regime.

II. WTO AND DOMESTIC REGULATORY AUTONOMY

A. WTO Objectives

The preliminary question that needs to be answered is what balance is struck between competing interests under the WTO agreements. The WTO system was, of course, intended to develop 'an integrated, more viable and durable multilateral trading system',[6] serving the GATT embodied goal of the 'substantial reduction of tariffs and other barriers to trade'.[7] However, while the system clearly promotes trade liberalization, it can be said that the WTO's core principle is non-discrimination.[8] This distinction goes to the very heart of the debate as the interpretation of 'necessity' is informed by the object and purpose of the treaties.[9]

In stark contrast to the harmonization/positive integration goals of other regimes such as the European Union (EU)[10] and the United States (US)

[4] The newer SPS, TBT and GATS agreements have not yet been subjected to the same degree of judicial scrutiny as they have only been raised and considered relevant on a limited number of occasions. As such, the discussion relating to these agreements is similarly constrained. In particular, while the GATS jurisprudence will be discussed, it will not be considered as a separate necessity test to the one contained in the GATT due to the almost identical textual content and similar goals of the agreements (see n 134). Necessity under the TRIPs Agreement will not be discussed.

[5] VCLT 1155 UNTS 331 art 31.

[6] WTO Agreement, preamble.

[7] Preamble to the GATT, echoed verbatim in WTO Agreement preamble.

[8] J O McGinnis and ML Movesian, 'Commentary: The World Trade Constitution', (2000) 114 Harv L Review 511, 517.

[9] VCLT art 31. This method of interpretation has been accepted by the adjudicating bodies on a number of occasions which have stated that this provision constitutes 'customary rules of interpretation of public international law' for the purposes of DSU art 3.2. See, for example, *United States-Section 301–310 of the Trade Act 1974*, WT/DS 152/R, 22 December 1999, paras 7.21–7.22; *United States-Gasoline*, WT/DS2/AB/R, 16; *US- Shrimp/Turtle*, WT/DS58/AB/R (1998) para 34.

[10] That is not to suggest that this is exactly what the EU has achieved: See T Mollers, 'The Role of Law in European Integration' (2000) 48 Am J Comp L 679, 683 noting it has achieved 'islands' of integration.

federal system,[11] which seek to create uniformity amongst their members in accordance with supra-nationally imposed standards, the WTO regime imposes no such requirements.[12] Instead, the WTO permits Members to implement regulatory and legislative regimes freely to promote whatever public policy objectives they deem to be in their national interests, with only one restriction: these measures cannot discriminate between imported and domestically produced goods of the same kind, or between trading partners.[13] Non-discrimination has been hailed for its facilitation of regulatory heterogeneity by identifying measures without excessive review of domestic policy choices.[14] However, in order to ensure sufficient protection for domestic measures designed to achieve non-trade goals, the Members included safeguards in the form of the article XX general exceptions. Under article XX, domestic policy choices aimed at protecting certain non-trade values are afforded such high importance that Members are permitted to escape their GATT obligations. Even measures that involve discrimination are acceptable but only if such discrimination is not arbitrary or unjustified.[15] This limitation on absolute freedom to regulate illustrates the dual objectives of article XX and embodies the broader challenge facing the WTO system as a result of its negative integration character.[16] Importantly, it also demonstrates that the resolution of competing interests in the WTO is the product of political negotiation.[17]

It is useful at this point to briefly define domestic regulatory autonomy since it is not defined in the WTO agreements. While subject to different interpretations,[18] the common understanding appears to be the freedom to pursue

[11] Note that Afilalo and Foster also refer to the NAFTA regime as requiring greater integration than the WTO: A Afilalo and S Foster, 'The World Trade Organisation's Anti-Discrimination Jurisprudence: Free Trade, National Sovereignty, and Environmental Health in the Balance' (2003) 15 Geo Int'l Envtl L Rev 633, 642–647.

[12] A number of exceptions to this general rule have developed including the obligations imposed under the TRIPS agreement which requires Members to achieve an agreed level of intellectual property protection within their jurisdictions (TRIPS Part II) while the SPS calls for SPS measures to be harmonized 'on as wide a basis as possible': (see SPS, art 3). Note, however, even these exceptions do not achieve complete harmonization as they only set minimum standards.

[13] GATT art I (Most-Favoured Nation) and art III (National Treatment).

[14] McGinnis and Movesian (n 8) 550.

[15] See art XX chapeau: '. . . nothing in this Agreement shall be construed to prevent the adoption or enforcement by any contracting party of measures . . .'

[16] The difficulties posed by the negative integration requirement are examined below. See also J R Pritchard and J Benedickson, 'Securing the Canadian Economic Union', in M. Trebilcock et al (eds), *Federalism and the Canadian Economic Union* (1983) ('[T]he tension between political autonomy and economic integration is inescapable in any non-unitary political system') 3.3; cited in J C Marwell, 'Trade and Morality: The WTO Public Morals Exceptions after Gambling' (2006) 81 NYUL Rev 802, 802.

[17] N Walker, 'The EU and the WTO: Constitutionalism in A New Key', in G De Burca and J Scott (eds), *The EU and WTO: Legal and Constitutional Issues* (2003) 47.

[18] Olivier Cattaneo notes that the concept of policy space is 'ambiguous' and used differently by different Members in different contexts, often in pursuit of conflicting objectives: 'Has the WTO Gone Too Far or Not Far Enough? Some Reflections on the Concept of Policy Space' in

domestic policy choices in accordance with the particular economic, social and cultural needs and/or preferences of that State.[19] Because the precise content of each term is difficult to discern (as seen by the WTO adjudicating bodies),[20] the manner of their interpretation can influence the trade/non-trade debate. For the purposes of this discussion, domestic regulatory autonomy will be defined as the freedom to pursue these goals, regardless of whether the value hierarchy accords with international preferences and regardless of its effect on trade.[21]

The concern to preserve regulatory autonomy while disciplining protectionist measures has characterized not only the evolution of the GATT but also the SPS and TBT agreements.[22] These latter two agreements were introduced to the world trade system as a direct response to the growing concern that internal measures could be and were being used as disguised restrictions on trade. In an attempt to minimize this practice, these agreements looked past non-discriminatory external trade barriers, required all measures to be the least trade-restrictive means available and introduced international standards. However, in keeping with the continued protection of domestic autonomy, the Members insisted that the establishment of such standards had to be negotiated in a manner that minimized the effects of regulatory actions on international trade without interfering in Members' responsibilities to legislate for the protection of their people and environment.[23] 'Necessity' was again used as the primary benchmark for testing the legitimacy of these internal measures.

Though the SPS and TBT agreements[24] were negotiated long after the GATT, the introduction of the concept of a 'single undertaking' creating a single treaty that must now be interpreted as a whole, ensures that the obligations contained in each agreement influence those contained in the others.

A. Mitchell (ed) *Challenges and Prospects for the WTO* (2005) 58; See generally, R Baldwin, C Scott, and C Hood (ed) *A Reader on Regulation* (1998).

[19] MM Du, 'Domestic Regulatory Autonomy under the TBT Agreement: From Non-Discrimination to Harmonisation' (2007) 6 Chinese JIL 269, 274.

[20] See in particular the difficulty faced by the adjudicating bodies in ascertaining what constituted public morals in the *US-Gambling* decision (Part II B).

[21] This definition adopts Du's perception of what regulatory autonomy means under the TBT Agreement (n. 19), however, the definition appears to fit equally well with what the Member's envisage beyond the TBT.

[22] G Marceau and J Trachtman, 'The Technical Barriers to Trade Agreement, the Sanitary and Phytosanitary Measures Agreement, and the General Agreement on Tariffs and Trade, A Map of the World Trade Organisation Law of Domestic Regulation of Goods' in G A Bermann and P C Mavroidis (eds), *Trade and Human Health and Safety* (2006) 9.

[23] Spec (71) 143, 30 September 1971, S III, art I(c) cited in Marceau and Trachtman, ibid 22.

[24] See above (n 1).

B. The Language of Necessity Contained in the GATT, SPS and TBT Agreements

1. The GATT

As mentioned, the GATT has provided for the pursuit of legitimate government policies that operate contrary to the general rules against discriminatory measures through article XX. Ten legitimate policy exceptions are listed and, though the list is exhaustive, the wording of each sub-clause is crafted broadly, giving wide scope for public policy measures to be covered by one of these exceptions. Of the ten exceptions, only three use the term 'necessary'.[25] The remaining exceptions require the measure to be 'relating to',[26] 'imposed for',[27] 'undertaken in pursuance of'[28] and 'essential to'.[29] Common to all ten, however, is the 'linking' character of each of these terms, that is, they measure the relationship between the ends sought and the means chosen. Each term requires a different degree of connection before the measure will legitimately fall under one of the designated exceptions. The very limited textual restrictions placed on domestic regulatory authorities to legislate with regard to these issues suggest that the negotiators intended for regulatory freedom to be preserved as much as possible and, concomitantly, to provide little opportunity for supranational review.[30]

Integral to the interpretation of the GATT necessity test but not an element of its immediate construction is its relationship with the chapeau of article XX. Referring to the introductory words of article XX, the chapeau prohibits measures falling within one of the ten exceptions from constituting: (i) arbitrary discrimination; (ii) unjustifiable discrimination; or (iii) a disguised restriction on international trade.[31] The purpose of the chapeau is recognized to be the prevention of the 'abuse' of the article XX exceptions by ensuring they are not 'so *applied* to frustrate or defeat the legal obligations of the holder of the right under the substantive rules of the General Agreement'.[32] The Appellate Body (AB) in *US-Gasoline*[33] affirmed the importance of the relationship between the article XX exceptions and the chapeau by proposing a two-tier test.[34] It noted that the contested measure not only needs to fall into

[25] GATT art XX (a), (b), (d). It is unclear whether (i) will be considered to refer to 'necessary to' or 'involving', the latter being preferred by the US-Gasoline AB, at 17. The sub-clause has not received any formal consideration and is worded differently from (a), (b) and (d).

[26] Sub-paras (c) and (e).

[27] Sub-para (f).

[28] Sub-para (h).

[29] Sub-para (j).

[30] Howse has suggested that the role of the trade insider network led to an 'amnesia' regarding the exact bargain that was struck between 'freer trade and the welfare state' and the development of 'an ideology of free trade': R Howse, 'The Boundaries of the WTO: From Politics to Technocracy—and Back Again: The Fate of the Multilateral Trading Regime' (2002) 96 AJIL 94, 98–99.

[31] All three requirements are deemed to 'impart meaning into each other' and 'can be read side by side': *US-Standards for Reformulated and Conventional Gasoline*, WT/DS2/AB/R, 29 April 1996 (*US-Gasoline*) 25.

[32] Emphasis added: ibid 22.

[33] ibid.

[34] ibid.

one of the ten exceptions to be regarded as justified protectionism but also to satisfy the introductory paragraph.[35]

Though the 'contours and content' of these three standards remain open,[36] these requirements have broader implications for the content and scope of the necessity test, given that a contested measure must satisfy both the exceptions provisions and the chapeau in order to succeed. The safeguards contained in the chapeau against the abuse of the exceptions provisions will logically limit what needs to be considered in the necessity test, especially given the breadth of the third prohibition.[37] However, the extent to which these three standards are sensitive enough to the particularities of a Member's regulatory environment remains unsettled.[38]

2. *The SPS and TBT agreements*

The interpretation of the SPS and TBT necessity tests has had a particularly strong influence over the development of the application of GATT article XX. However, there are number of significant differences between each test. In particular, unlike the GATT where the necessity provisions are phrased as exceptions to Members' obligations, the SPS and TBT necessity requirements impose positive obligations, shifting the burden of proof from the defending State to the complaining State to show the measures were not necessary. The textual content of each test thus logically differs.

At first glance, the agreements appear similar. The SPS preamble echoes the chapeau of article XX and sub-clause (b), indicating from the outset that the purpose of the agreement is to achieve a balance between allowing Members sufficient policy space to regulate with regard to SPS measures while disciplining protectionist measures. The TBT preamble also contains a very similar paragraph with an additional paragraph affirming the Members' desire to ensure that such regulations do not constitute 'unnecessary obstacles to international trade'.[39] However, after this point, the language of the necessity provisions differs from that contained in the GATT.

[35] ibid, citing *US-Imports of Certain Automotive Assemblies* Report adopted 26 May 1982 BISD 30S/107 para 56; *Argentina-Measures Affecting the Export of Bovine Hides and the Import of Finished Leather* WT/DS155/R, 19 December 2000, paras 11.288–11.289.

[36] *US-Import Prohibition of Certain Shrimp and Shrimp Products*, WT/DS58/AB/R, 12 October 1998 (*US-Shrimp/Turtle*) para 120: After stating that the standards were 'necessarily broad in scope and reach', the AB stated the scope of these three standards 'will vary as the kind of measure under examination varies'.

[37] TJ Schoenbaum, 'International Trade and Protection of the Environment: The Continuing Search for Reconciliation' (1997) 91 AJIL 269. Schoenbaum argues in favour of greater deference from the adjudicating body to the regulatory autonomy of the State when applying the exception provisions under art XX as the next logical step is to apply the chapeau: 277.

[38] For a fuller exploration of the potential for the chapeau to accommodate non-trade issues see S Gaines, 'The WTO's Reading of the GATT art XX Chapeau: A Disguised Restriction on Environmental Measures' (2001) 22 U Pa J Int'l Econ L 739.

[39] TBT preamble, paras 5 and 6.

Article 2.2 of the TBT agreement requires technical regulations to be 'not more trade-restrictive than necessary to fulfil a legitimate objective', thus automatically importing a least-restrictive means test which is not textually present in article XX. It also includes a non-exhaustive list of legitimate objectives of similar character to the exceptions contained in the GATT. Article 2.2 of the SPS requires that the SPS measure is applied 'only to the extent necessary to protect human, or plant life or health' while SPS article 5.6 includes a least restrictive means test similar to the one employed under the TBT. It requires SPS measures to be 'not more trade-restrictive than required to achieve their appropriate level of . . . protection, taking into account technical and economic feasibility'. The provision is accompanied by a footnote that adds a further qualification to the implementation of the test, that is, the reasonable availability of the alternative measures.[40]

The significance of these distinctions will be explored in Part II, though even at this preliminary stage it appears that the balance struck between free trade and regulatory autonomy under each agreement appears to be slightly different. Interestingly, the extent to which the jurisprudence has given effect to these distinctions is open to question.

C. *Interpretative Tools and the Horizontal/Vertical Power Struggle*

The inextricable relationship between a Member's choice of measure and its justification makes adjudicating its legitimacy impossible without some form of intrusion into the decision-making authority of the State. The issue that remains unresolved is how much interference by the adjudicating bodies should be permitted and in what manner. Not only has the line been repeatedly drawn, shifted and erased in academic discourse[41] (exacerbated by the growing diversity of WTO membership),[42] the jurisprudence has echoed this uncertainty in its application of the necessity test, reaffirming the link between the horizontal trade/non-trade debate and the vertical institutional struggle between Members and the adjudicating bodies.

The relatively low level of integration required under the WTO treaty makes the appropriate delineation of competence all the more difficult to define. In contrast, the ultimate vision of a unified community and single market has provided the EU with a significant advantage over the WTO in this regard, essentially providing the European Court of Justice (ECJ) with the clear constitutional prerogative to curtail state sovereignty in favour of free trade and thus the greater economic welfare of all nations within the community.[43]

[40] SPS Agreement, art 5.6, fn 3.
[41] McGinnis and Movesian (n 9) 550.
[42] See Marwell (n 16) 808–809.
[43] Button makes this comment specifically in relation to health measures adopted by states: above (n 2) 208–9.

Even within this highly integrated community, ECJ practice has demonstrated a deferential tendency, only striking down measures in extreme cases.[44] In addition, in cases where there is a paucity of facts in preliminary rulings, the Court will leave the final decision on the proportionality of the measure to the national courts.[45]

This commonality has also helped define the level of scrutiny in the US federal structure in instances where state measures have implicated inter-state commerce. The US Commerce Clause[46] has been used by the Supreme Court to subject the measure to intense scrutiny in order to protect the aims of the federal structure that embodies the idea that the future of each State is inextricably linked to the fate of the others.[47] Unlike the EU and US, the WTO Members have no common history or common cultural values upon which it could build a community. Nor, most importantly, is this its goal. As can be seen, the character of a negative integration treaty such as this one necessarily implies that in addition to the stated goal of trade liberalization and non-discrimination, a respect for sovereignty and national deference is also prominently present.[48] Additionally, nothing in the text of the treaty indicates that national sovereignty should be given lesser weight than the other principles. In fact, it is this continuing respect for such autonomy that characterises the WTO as a negative integration treaty.

Contributing to the uncertainty is the fact that a principled legal doctrine on the standard of review[49] in the WTO context has not yet been formulated.[50] The treaty text is devoid of any explicit reference to what it should be.[51] However, that is not to say that the treaty is completely silent on the matter: as

[44] Joanne Scott argues that the ECJ has been more deferential to the political imperatives driving the introduction of certain measures by the national government: J Scott, 'On Kith and Kine (and Crustaceans): Trade and Environment in the EU and the WTO' in JHH Weiler (ed), *The EU and WTO and the NAFTA*, *Towards a Common Law of International Trade* (2000) 125–68. cf Button who argues that the ECJ case law shows 'very little deference' to Member State preferences when applying the least trade-restrictive alternative test: ibid 208.

[45] J Neumann and E Turk, 'Necessity Revisited: Proportionality in World Trade Organization Law after *Korea-Beef, EC-Asbestos*, and *EC-Sardines*' (2003) 37 J World Trade 1 199, 232.

[46] Art 1, s 8, clause 3 of the US Constitution states that Congress has the exclusive authority to manage commerce between the states, with foreign nations and Indian tribes.

[47] Button (n 2) 202.

[48] Hilf has highlighted eight principles and objectives of the WTO found in prominent positions in the legal text and acknowledged by the AB.: M Hilf, 'Power, Rules and Principles—Which Orientation for WTO/GATT Law?' (2001) 4 J Int'l Econ L 111.

[49] The term 'standard of review' in this discussion is used in its most general form to refer to its influence over the general allocation of power between the adjudicating bodies and WTO Members by dictating the extent of power to be awarded to each party on issue of law and fact: M Andenas and S Zleptnig, 'Proportionality: WTO Law in Comparative Perspective' (2007) 42 Tex Int'l L J 371, 395. Consequently, the discussion relating to the application of the 'necessity' test will involve reference to both types of review. See also M Oesch, *Standards of Review in WTO Dispute Resolution* (OUP, Oxford, 2003) for greater consideration of this issue.

[50] Button (n 2) 193.

[51] With the exception of art 17.6 of the *Agreement on Implementation of Article VI of the GATT*.

will be shown, much can be inferred from the particular drafting of certain provisions, particularly those relating to necessity under the different agreements. However, the clearest pronouncement on the issue to date came from the AB in *EC-Hormones*[52] where it decided that the general standard of review to be applied to all WTO agreements was simply 'an objective assessment of the matter'.[53] It believed such a standard reflects the balance 'between jurisdictional competences conceded by the Members to the WTO and the jurisdictional competences retained by the Members for themselves'.[54] The ambiguity of this standard was little assisted by the AB's outright rejection of the two extreme possibilities: de novo and total deference standards.[55] The result has been the application of a different standard of review under each WTO agreement.[56]

In light of the above, it is not surprising that the extent to which the interpretation of necessity has restricted the autonomy of WTO Members has varied significantly. The changes to the test have occurred through the adjudicating bodies' use of what Trachtman terms 'trade-off' devices, that is, the doctrinal tools used to balance competing interests.[57] Each of these devices infers a different standard of review by informing the vertical allocation of powers between adjudicator and Member as well as the horizontal conflict between trade liberalization and other policy goals.[58] At one end of the scale, simple means-ends[59] rationality tests require little more than a face-value examination of the State's regulatory choices in order to determine whether the measure is directed at the goal the State claims it is. The traditional least-restrictive means test takes a step beyond the ends-means rationality test and asks whether the State chose the most trade-efficient option available to it to achieve its goal.[60] Thus, the characterization of the end sought will often be determinative.[61] The reasoning provides more scope for the adjudicating body to review the measure through the determination of what measures are reasonably available and whether the purpose can be legitimately pursued.[62]

Closer to the opposite pole, proportionality, balancing and cost-benefit analysis are trade-off devices that provide far more opportunity for judicial

[52] *EC-Measures Concerning Meat and Meat Products (EC-Hormones)*, WT/DS48/AB/R, 16 January 1998, para 116.

[53] ibid: The AB based its reasoning on art 11 of the *Understanding on Rules and Procedures Governing the Settlement of Disputes, WTO Agreement, Annex 2 (DSU)*.

[54] ibid, para 115.

[55] ibid, para 117.

[56] Andenas and Zleptnig (n 49) 396.

[57] Trachtman identifies six broad categories: national treatment, simple means-end rationality test, necessity or least trade restrictive alternative test, proportionality test, balancing test and cost-benefit analysis: see JP Trachtman, '*Trade and . . . Problems, Cost-Benefit Analysis and Subsidiarity*', available at: http://ejil.org/journal/Vol19/No1/art3.html, 1.

[58] ibid 2.

[59] Each of these terms should be interpreted in their most general sense.

[60] Trachtman (n 57) 3.

[61] ibid.

[62] ibid.

intervention into the decision-making processes of States. Making a number of appearances in the application of the necessity test in recent years,[63] proportionality *strictu sensu* has been termed a 'European law' concept[64] and requires an examination of whether the means adopted are proportionate to the ends sought. In the WTO context, it would mean that the restriction of trade must not be disproportionate to the benefits arising from the protection of the value covered by the measure. In this respect, it is largely similar to cost-benefit analysis but lowers the very high hurdle that analysis imposes by not necessarily requiring that the benefits outweigh the costs.[65] Importantly, it also requires a comparative analysis of the outcomes of the proportionality test as applied to alternative measures.[66] Thus, this device parallels the elements of the necessity test but places the value being pursued under scrutiny as well as the means chosen to achieve it.

Finally, the language of the balancing test has been used frequently in WTO necessity interpretations.[67] On a continuum from least intrusive to most intrusive into regulatory autonomy, this device can be seen as inferring a more intrusive standard of review than the one contained in the necessity test but arguably less intrusive than proportionality *strictu sensu* and the cost-benefit analysis. This is because it does not actually attempt to quantify the competing values but rather 'recognizes the difficulty in formalizing the analysis' and instead seeks to balance all the competing values in a less stringent manner.[68] However, again, the test requires an evaluation of the value being pursued.

The issue of contention thus arising from the use of the latter three devices when applying the necessity test is whether and to what extent the unelected and unaccountable adjudicating body can and/or should engage in an exercise of 'second-guessing' the policy choices of popularly elected national governments.[69] The absence of a clear standard of review has allowed the delineation of competencies to fall, by default, into the hands of the adjudicating bodies. However, that is not to say the treaty is silent on the issue: as will be demonstrated, the specific language chosen to construct each necessity test under each agreement provides a solid groundwork for the resolution of this issue. Thus, the fundamental question is whether the scope afforded to

[63] Neuman and Turk have concluded that the WTO tribunals have not yet adopted a strict proportionality test and that the rules of the WTO necessity test do not incorporate any explicit reference to it: Neuman and Turk, above (n 45) 231. In contrast, Hilf regards it as 'one of the more basic principles underlying the multilateral trading system': Hilf, above (n 48) 6. Desmedt concluded in his analysis that there is no uniform interpretation of the proportionality principle in WTO law: A Desmedt, 'Proportionality in WTO Law' (2001) 4 (3) J Int'l Econ Law 441, available at http://jiel.oxfordjournals.org/cgi/content/abstract/4/3/441, 21.

[64] M Kennett, J Neuman and E Turk, 'Second Guessing National Level Policy Choices: Necessity, Proportionality and Balance in the WTO Services Negotiations' (2003) CIEL, presented at the WTO's 5th Ministerial Meeting in August 2003, available at: http://www.ciel.org/Publications/Necessity_3Sep03.pdf at 5.

[65] Trachtman (n 57) 3.

[66] ibid.

[67] See Part IIB.

[68] Trachtman (n 41) 3.

[69] McGinnis and Movesian (n 8) 513.

domestic autonomy by the adjudicating bodies in both the horizontal and vertical struggles accords with the level of interference negotiated and agreed upon by the Members. It is to this issue we now turn.

III. A CRITIQUE OF THE NECESSITY TEST JURISPRUDENCE

A. The Necessity Test Under the GATT Article XX

1. The original test: 'least restrictive means' and 'reasonable availability'

The term 'necessary' was first interpreted[70] by the GATT Panel in *US- Section 337 of the Tariff Act of 1930*[71] when the US claimed that the measure in question was necessary under GATT article XX(d) to secure compliance with domestic patent laws. The US argued that section 337 provided the only means of enforcement of United States patent rights against imports of products manufactured abroad by means of a process patented in the United States. Here the Panel stated:

> [a] contracting party cannot justify a measure inconsistent with another GATT provision as 'necessary' if an alternative measure which it could reasonably be expected to employ and which is not inconsistent with other GATT provisions is available to it.[72]

The Panel went on to state that in cases where no GATT-consistent measure is reasonably available, 'a contracting party is bound to use, among the measures reasonably available to it, that which entails the least degree of inconsistency with other GATT provisions'.[73] In determining whether such alternatives existed, the Panel cited the existence of other possible measures without regard to whether these alternatives could achieve the level of protection desired by the US,[74] and despite US arguments that they could not.[75] Significantly, the Panel stated that the LRM requirement did not require a change in either the Member's chosen level of enforcement or the substance of the law, provided that the law did not distinguish between domestically produced and imported goods.[76] This suggests that it was aware that its choice of judicial tool could be seen as an incursion into the regulatory freedom of WTO Members. As the application of the test showed that this extra-textual obligation requires Members to assess the alternative only against its relative impact on trade,

[70] Although *US-Import of Certain Automotive Spring Assemblies*, BISD 30S/107 adopted 26 May 1983, raised the issue, it was never considered by the panel.

[71] *US- Section 337 of the Tariff Act of 1930*, BISD 365/345 adopted 7 November 1989 (*US- Section 337*).

[72] *US- Section 337*, BISD 365/345, para 5.26; Reaffirmed in *US-Measures Affecting Alcoholic and Malt Beverages* (*US- Malt Beverages*) adopted 19 June 1992, BISD 39S/206, para 5.52; and *Thailand-Restrictions on Importation of and Internal Taxes on Cigarettes* (*Thai-Cigarettes*) adopted 7 November 1990 BISD 37S/200, para 223.

[73] *US-Section 337*, BISD 365/345, para 5.26.

[74] ibid para 5.33. [75] ibid para 4.9. [76] ibid para 6.1.

it is difficult to see how the LRM necessity test could peacefully co-exist with the Member's right to choose their level of protection. The issue of whether a Member's treaty-based freedom to pursue a legitimate policy goal could be maintained under the LRM test without contradiction was now raised.[77]

This particular application of the LRM test in the *Thai-Cigarettes*[78] case did little to remedy this concern. In this case, Thailand argued that the import prohibition it had instituted on tobacco and tobacco products while simultaneously allowing the sale of domestic cigarettes was necessary 'to protect human . . . life or health' under article XX(b). The Panel noted that while article XX(b) 'clearly allowed contracting parties to give priority to human health over trade liberalization', and accepted this to be Thailand's goal,[79] any measures implemented to achieve this goal could only be considered necessary if 'there were no alternative measures consistent with the General Agreement, or less inconsistent with it, which Thailand could reasonably be expected to employ to achieve its health policy objectives'.[80] Thus, although *US-Section 337* had dealt with the meaning of 'necessary' under article XX(d), the Panel in *Thai Cigarettes* concluded that the term 'necessary' had the same meaning under sub-clause (b) and (d).[81] The decision suggested that a uniform interpretation was being developed with regard to necessity under GATT article XX, despite the different values being pursued under each exception. While this approach finds support in the fact that article XX does not purport to place the values being protected in any hierarchical order, the different wording of the two provisions, as well as the fact that the objects and values of the 'laws and regulations' in sub-paragraph (d) are undefined while the values contained in (b) are clearly expressed, may constitute grounds for questioning the appropriateness of this uniformity.

Following an examination to this effect, the Panel found that alternatives existed with regard to both of Thailand's qualitative and quantitative aims.[82] Particularly significant was the Panel's acceptance of the US-proposed alternative to reduce demand: a ban on the advertisement of both domestic and foreign cigarettes, despite Thailand's arguments and evidence suggesting that

[77] See generally, D Regan, 'The Meaning of "Necessary" in GATT Article XX and GATS Article XIV: The Myth of Cost-Benefit Balancing' (2007) 6 World Trade Review 3 347, 348, whose analyses this issue with regard to the balancing test applied in *Korea-Beef*, 348.

[78] *Thai-Cigarettes* DS10/R- 37S/200, adopted on 7 November 1990.

[79] ibid para 76.

[80] ibid para 75.

[81] The Panel concluded, 'In both paragraphs the same term was used and the same objective intended: to allow contracting parties to impose trade restrictive measures inconsistent with the General Agreement to pursue overriding public policy goals to the extent that such inconsistencies were unavoidable. The fact that paragraph (d) applies to inconsistencies resulting from the enforcement of GATT-consistent laws and regulations while paragraph (b) applies to those resulting from health-related policies therefore did not justify a different interpretation of the term "necessary"': *Thai-Cigarettes* DS10/R- 37S/200, para 74.

[82] ibid paras 77–81.

such measures were ineffective to achieve its stated aims.[83] Further, no consideration was given to whether these alternatives were feasible with regard to Thailand's particular social, political and economic conditions.

The effect of the application of the LRM necessity test in this manner has two points of significance for the balance of domestic regulatory autonomy against free trade. First, as also seen in the *US-Section 337* Panel decision, the *Thai-Cigarettes* Panel indirectly engaged in judging the value of the policy goal and undermined the Member's right to choose its own level of protection by deciding in favour of the alternatives without enquiring into their ability to achieve Thailand's stated goals. The implication of this reasoning is to limit a Member's freedom to pursue non-trade goals if they negatively affect trade to a degree which is determined to be unacceptable by a Panel. Secondly, and related to the first issue, the Panel's decision has fallen subject to intense criticism for its insensitivity to the practical regulatory experience of governments.[84] Importantly, the decision reveals an absence of any substantive consideration of the term 'reasonably available' and how it should be applied. All that was required to defeat the claim of necessity was the existence of a hypothetically available alternative. Thus, the application of the LRM test raises a number of concerns relating not only to the appropriateness but also the competence of the adjudicating bodies to assess the feasibility of regulatory alternatives over the decision made by a Member's government that possesses a greater understanding and knowledge of the internal particularities of the State.[85]

The question of reasonable availability was given similar treatment in *US-Gasoline*.[86] Again, the Panel found that reasonable alternatives existed, despite the US's claims and evidence that such alternatives would be less effective, ineffective or unfeasible.[87] Of particular interest is the Panel's finding that the determination of the gasoline's origin (an issue of particular concern to the US) 'would *often* be feasible',[88] thereby implicitly lowering the higher level of protection the US had sought to achieve. Finally, the Panel

[83] Thailand relied on the World Health Organisation's findings that advertising bans were circumvented by multinational tobacco companies through indirect advertising and other modern marketing techniques, ibid, paras 27 (Thailand's submissions), para 55 (WHO submissions).

[84] D Osiro, 'GATT/WTO Necessity Analysis: Evolutionary Interpretation and its Impact on the Autonomy of Domestic Regulation' (2002) 29 Legal Issues of Economic Integration 2, 123, 127–8; M J Trebilcock and Robert Howse, *The Regulation of International Trade* (3rd edn, 2005) 518; Neumann and Turk (n 45) 208.

[85] Button (n 2) 30. See generally, C Correa, 'Implementing National Public Health Policies in the Framework of the WTO Agreements' (2000) 34 J World Trade 5, 89.

[86] *US-Standards for Reformulated and Conventional Gasoline*, WT/DS2/R, 29 January 1996 (*US-Gasoline*).

[87] See for example, the US argument regarding the difficulty of exercising enforcement jurisdiction with respect to a foreign refinery, the 'impossibility of determining the refinery origin for each imported shipment', and the recognition of the fact that not all refineries were able to produce the evidence required to establish an individual baseline: *US-Gasoline* WT/DS2/R, para 6.23.

[88] ibid para 6.36 (emphasis added).

stated that an alternative measure did not cease to be reasonably available simply by virtue of administrative difficulties.[89] Thus, while the term 'reasonably available' may appear, on its face, to provide greater scope for accommodating the particular social and economic demands of each state, its interpretation by the Panel in this case nullified this potential.

2. *Textual consistency of the 'least restrictive means'*

The application of the test in this way also shows that two possible versions of the LRM test exist: first, a strict analysis where the measure is deemed invalid only if the available alternatives achieve exactly the same benefits at a lower cost to free trade; and secondly, a more lax test under which the adjudicating body may sacrifice some benefits provided by the contested measure in order to implement the alternative 'where those benefits are judged by the court to be of less value than the cost–avoidance achieved by the change'.[90] By favouring the latter option, the necessity test began to closely resemble a cost-benefit analysis under which the adjudicating bodies inserted their own value judgments with little deference afforded to the factual findings of the defending states. This is especially so, as seen in both the *Thai-Cigarettes* and *US-Gasoline* decisions, where the adjudicating bodies dismissed the arguments of the responding Member as to whether the alternatives were 'reasonably available'.

When compared with the text of article XX and the object of the GATT as a whole, this line of cases illustrates that the LRM test was used to tilt the balance in favour of trade liberalization beyond that provided in the treaty. The text of each necessity exception requires the proposed measure to be 'necessary to' achieve a stated goal. Despite affirming that the adjudicating body's task is to 'address whether the inconsistent measure was necessary to achieve the policy goal sought',[91] in practice, it has been used to refer to the justifiable extent of inconsistency between the regulation and the obligations under the treaty instead of the measure's relationship with fulfilling the goal being sought. As Schoenbaum has noted, 'necessary no longer relates to the protection of living things, but to whether or not the measure is a "necessary" departure from the trade agreement.'[92] Though his criticism was directed to

[89] ibid para 6.28. The Panel's finding relating to its application of the necessity test under arts XX (b) and (d) were not appealed so the issue did not receive fuller treatment.

[90] Regan makes this distinction, calling the latter test the 'loose LRA (least-restrictive alternative) test' and notes that the existence of alternatives that achieve all the same benefits as the contested measure are rare: Regan, 'Judicial Review of Member-State Regulation of Trade within a Federal or Quasi-Federal System: Protectionism and Balancing, Da Capo' (2001) 99 Mich L Rev 1853, 1899–1900; See also Sykes for the argument that the LRM test is being applied as a form of 'crude cost-benefit analysis': A Sykes, 'The Least Restrictive Means' (2003) 70 U Chi L Rev 403.

[91] *US-Gasoline* WT/DS2/R, para 6.22.

[92] Schoenbaum, above (n 37) 276. See also R J McLaughlin, 'Sovereignty, Utility and Fairness: Using US Takings Law to Guide the Evolving Utilitarian Balancing Approach to Global

the interpretation of article XX(b), his comments can be extended to all three of the article XX exceptions referring to necessity. The distinction is further supported by the construction of the necessity tests in the texts of the TBT and SPS which clearly provide for the introduction of some LRM analysis, highlighting its absence in the text of the GATT.[93]

In further support for this view, Schoenbaum argues that this new interpretation amounts to an odd distinction between those provisions using the words 'necessary to' and the remaining sub-clauses.[94] He states that to allow the interpretation of necessity to include the LRM test applied by the adjudicating bodies results in a higher threshold for the survival of measures regulating the protection of human life and health than those 'undertaken in pursuance of obligations under any intergovernmental commodity agreement'.[95] However, it appears that this criticism does not take account of the fact that the text of the treaty appears to make precisely this distinction by requiring measures purporting to regulate health and public morals to be 'necessary' rather than simply 'relating to', 'involving' or 'imposed for' which all plainly require less of a connection between the measure and the goal than the term 'necessary'. While the reasons for this drafting distinction are not clear, it may be that the drafters foresaw the potential abuse of these provisions given the political sensitivity and malleable character of these goals in particular and thus sought to restrict their use to instances where this higher threshold could be satisfied. The distinction further highlights the need for the interpretation of necessity to give effect to these complex considerations and shows that balancing the competing objectives under the GATT is no simple task.

Attention must also be paid to the function of the necessity test within the agreement as a whole. The LRM test as applied by the adjudicating bodies appears to ignore the character of the article XX sub-clauses as *exceptions* to the obligations contained in the agreement. The extent to which such measures are inconsistent with the GATT should thus be irrelevant. Further, the introduction of the concept of 'less-GATT inconsistent' measures is curious in itself, not only because it is difficult to see how this can be measured,[96] but also because the meaning of the word does not provide space for shading.[97]

This interpretation has important ramifications for the continued relevance of the chapeau, which is specifically worded to examine the extent to which the discriminatory effects are justifiable, thus also contradicting the two-step

Environmental Disputes in the WTO' (1999) 78(4) Oregon Law Review 855, 889 where he argues that Panels had traditionally relied on 'strained textual and functional definitions to prevent article XX exemptions from being applied'.

[93] See Part IIB below.

[94] Schoenbaum (n 37) 276.

[95] GATT art XX(h).

[96] An issue that has not been given any attention by the adjudicating bodies.

[97] Camilleri also notes this latter point: V Camilleri, 'An Analysis of the Necessity Tests as Applied by the WTO adjudicator' (LLM Dissertation, College of Europe University, 2004) 9.

analysis provided for by the AB in *US-Gasoline*.[98] By failing to give effect to all the terms of the treaty, the application of this particular LRM test appears to violate VCLT article 31,[99] as interpreted by the Appellate Body itself.[100]

B. Reform: The 'Weighing and Balancing' of Necessity

The Appellate Body's decision in *Korea-Beef*[101] has been generally recognized as marking the beginning of a new interpretation of the meaning of necessity and its application under article XX through the introduction of a cost-benefit balancing test.[102] The test was further refined and reinforced in *EC-Asbestos*.[103] However, the extent to which the *Korea-Beef* balancing test (hereinafter *Korea-Beef* test) actually changed the nature of the existing LRM test or simply confirmed and refined the content of the weighing process implicitly present in the application of that test is open to question. Either way, the test has had broad ramifications for the trade/non-trade debate by affording greater power to the adjudicating bodies to balance the effects of the regulatory goal.

1. The new 'necessary' under the GATT

In *Korea-Beef*, Korea unsuccessfully claimed that its dual retail system segregating imported and domestic produced beef was justified under article XX(d), alleging that such differential treatment was necessary to protect consumers against fraudulent practices prohibited under its Unfair Competition Act.[104] After accepting that Korea's measures were covered by the policy objectives envisaged in article XX(d), the AB began its application of the necessity test. First, the AB noted that 'necessary' does not always mean 'indispensable' or 'inevitable' but refers instead to 'a range of degrees of necessity'.[105] While one extreme of the continuum defines necessity as 'making a contribution to', the AB stressed that a necessary measure in the context of the GATT is 'located significantly closer to the pole of indispensable'.[106] In this respect, the AB appeared to be paying closer attention to the textual provisions of the treaty and giving full effect to the distinctions of each sub-clause.

[98] See discussion of chapeau (n 34) and surrounding text.
[99] Schoenbaum (n 37) 277.
[100] *US-Gasoline*, WT/DS2/AB/R, 23.
[101] *Korea-Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161.169/AB/R, 11 December 2000, (*Korea-Beef*).
[102] Note, however, Regan's argument that this is a misunderstanding of what the test actually entails: Regan (n 77).
[103] *EC-Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, 12 March 2001, (*EC-Asbestos*).
[104] It alleged that such differential treatment was necessary to protect consumers against fraudulent practices prohibited under its Unfair Competition Act.
[105] *Korea-Beef*, WT/DS161.169/AB/R, para 161.
[106] ibid.

In some respects, the AB appeared to continue this trend, relying heavily on the context of sub-clause (d) to assert that, as it was 'susceptible to being applied in respect of a wide variety of laws and regulations,' a treaty interpreter 'may, in appropriate cases, take into account the relative importance of the common interests or values that the law or regulation to be enforced is intended to protect.'[107]

Thus, the AB reasoned that a finding of necessity under article XX(d) required 'a process of weighing and balancing' at least three factors:

> . . .which prominently include the contribution made by the compliance measure to the enforcement of the law or regulation at issue, the importance of the common interests or values protected by that law or regulation, and the accompanying impact of the law or regulation on imports or exports.[108]

In implementing this balancing test, the AB noted, '[t]he more vital or important those common interests or values are' and the greater its contribution, 'the more easily the measure might be considered "necessary"'.[109] Finally, a measure with a 'relatively slight impact upon imported products might more easily be considered "necessary" than a measure with intense or broader restrictive effects.'[110] Importantly, the AB's wording leaves open the list of factors that can potentially be taken into account when engaging in the balancing process but does not provide any guidance on how these factors interact.

Though many issues regarding the test's application were unclear, it appears that examination of the measure in question is, at first, self-contained: that is, the weighing and balancing proceeds without reference to the relative trade restrictiveness of alternative measures.[111] If, following the application of this process, the contested measure can prove itself to be reasonably effective in protecting an important interest with moderate trade restrictions, all alternative measures involving less GATT-inconsistent restrictions will be considered not 'reasonably available' automatically.[112] However, if weighing the contested measure does not produce this outcome, the next step will be to examine the availability of alternative measures in relative terms to the contested measure.[113] Thus, a distinction was made between indispensable and dispensable measures. The problem of course with maintaining this apparent bifurcation in practice is that it is difficult to conceive of any situation where a contested measure that imposes trade restrictions to any degree will be deemed 'necessary' without entering into a consideration of whether reasonable less-trade restrictive alternatives exist.

With regard to the issue of reasonable availability, the AB implicitly adopted the Panel's statement that the onus was on the responding state to

[107] ibid para 162.
[108] ibid para 164.
[109] ibid para 163.
[110] ibid.
[111] Neumann and Turk (n 45) 211.
[112] ibid.
[113] *Korea-Beef*, WT/DS161.169/AB/R, para 173.

show that the alternatives were not reasonably available or 'unreasonably burdensome' in an economic or technical sense, 'taking into account a variety of factors including the domestic costs of such alternative[s]'.[114] However, no indication was given as to how this would be judged, and what degree of deference (if any) would be given to a Member's assessment of its own circumstances. The AB's decision that Korea could achieve its desired goal using WTO-consistent measures 'if Korea would devote more resources to its enforcement efforts on the beef sector',[115] despite Korea's claims that sufficient resources were not available, highlight the significance of these questions for the protection of Members' regulatory choices. It also raises concern about the adjudicating bodies' competence to make such a determination with limited knowledge of the particularities of each Member's economy.

The AB's recognition of Korea's autonomy over its chosen level of protection was also called into question. Contrary to Korea's claims, the AB decided that Korea had not in fact intended to totally eliminate fraud but rather wished only to 'reduce [the opportunity for fraud] considerably'.[116] It arrived at this conclusion solely on the basis that the total elimination of fraud would 'probably require a total ban of imports'.[117] As a consequence of this finding, the alternatives to be considered only had to meet this lower level of protection inferred by the AB.

The AB concluded that Korea had failed to show that GATT-consistent alternative measures were not reasonably available to it and thus the measure failed under article XX(d). Significantly, despite setting out the balancing test, the AB did not engage in applying the test to the facts of the case.[118] Its conclusion appears to suggest that survival of a contested measure under the LRM test as it was previously applied was still the main benchmark, thought now (at least theoretically) supplemented with a proportionality test if not a cost-benefit analysis. The main distinction was that the value of the goal pursued was expressly called into question.

The *EC-Asbestos* AB contributed significantly to the development of this new necessity testthrough the special attention paid to the concept of 'reasonably available'. The appeal involved an examination of whether a French ban on the manufacturing, sale and import of asbestos fibres introduced to protect human life and health was 'necessary' under article XX(b). No justification was given for the transfer of the necessity test under article XX(d) to (b).[119] First, it removed the untenable bifurcation and

[114] ibid.
[115] ibid para 180.
[116] ibid para 178.
[117] ibid.
[118] The Panel, however, did specifically state that the dual retail system was 'a disproportionate measure not necessary to secure compliance with the Korean law against deceptive practices': Panel Report para 675.
[119] The AB merely stated that the necessity test formulated in section 337 as modified by *Korea-Beef* was correctly applied to the case by the Panel: *EC-Asbestos*, WT/DS135/AB/R, paras 78–79. It is interesting to note that that AB was not obliged engage with the art XX analysis as it decided against the Panel's findings that the art III:4 had been violated. Neumann and Turk

affirmed the *Korea-Beef* test as an integral part of the determination of whether the policy goal could be achieved by a reasonably available less trade-restrictive alternative.[120] This approach was confirmed again in the most recent application of the test: *Brazil-Tyres*.[121] Here the AB stated that even if a consideration of the relevant factors produces a preliminary conclusion that the measure is necessary, 'this must be confirmed by comparing the measure with its possible alternatives.'[122]

The second point of clarification by the *EC-Asbestos* AB was that reasonableness now meant that a measure with less impact on trade could not be considered a reasonable alternative if it could not achieve the same degree of protection sought,[123] thereby reducing the impact of the 'trade trumps' ideology displayed in the preceding cases. The AB reaffirmed that every Member had an 'undisputed right to choose the level of protection they consider appropriate in a given situation'.[124] Arguably, the extent of deference afforded to the regulatory autonomy of the State by this AB was greater than that afforded to Korea as the AB held that a State is not bound to automatically follow the prevailing majority scientific opinion when drafting its health policy.[125] Further, the AB accepted that France had chosen to 'halt' the spread of asbestos-related health risks by structuring the measure to eliminate it completely,[126] despite the existence of a number of small exceptions to the ban. The AB used the level of protection chosen by France as its primary benchmark for determining whether Canada's suggested alternatives were reasonably available.[127] If the logic of the AB in *Korea-Beef* is applied to this factual scenario, the regulatory exceptions could have been used by the AB to negate the argument that France sought to totally eliminate asbestos products but had instead sought only to reduce it considerably.[128] The difference between findings in the two cases highlights the extent to which the preservation of domestic regulatory autonomy lies in the discretionary hands of the adjudicating bodies under the *Korea-Beef* test.

Finally, the AB declined to address the issue of whether administrative difficulties in the implementation of the measure would impede it from being considered reasonably available, thus leaving open the question of how 'difficulty' will be assessed.[129]

However, the decision did not clarify how the elements of the balancing test would interact. In applying the test the AB did not refer to the

have interpreted the AB's decision to continue with the art XX analysis as a reflection of the AB's eagerness to 'make judicial policy': Neumann and Turk (n 45) 213.

[120] *EC-Asbestos*, WT/DS135/AB/R, para 172.

[121] *Brazil-Measures Affecting Imports of Retreaded Tyres*, WT/DS332/AB/R, 3 December 2007.

[122] ibid para 156.

[123] *EC-Asbestos*, WT/DS135/AB/R, para 174.

[124] ibid para 168.

[125] ibid para 178.

[126] ibid paras 172–4.

[127] ibid para 174.

[128] See (n 115) and surrounding text.

[129] ibid; *EC-Asbestos*, WT/DS135/AB/R, para 169.

trade-restrictiveness of the measure at all and the two remaining elements were looked at in isolation without reference to how they affected each other.[130] It found the measure to be effective in meeting the desired goal by reference to the fact that it was a total ban and referred only to importance of the value, stating:

> In this case, the objective pursued by the measure is the preservation of human life and health through the elimination, or reduction, of the well-known, and life-threatening, health risks posed by asbestos fibres. The value pursued is both vital and important in the highest degree. The remaining question, then, is whether there is an alternative measure that would achieve the same end and that is less restrictive of trade than a prohibition.[131]

It appears, then, that the 'necessity' of a measure under the GATT is found by default after consideration of the alternatives. Further, while this statement did reinforce the fact that the weighing and balancing test would not involve balancing the level of protection against the trade restriction[132] (as appeared in the earlier applications of the LRM test), it highlighted the extent to which the importance of the value being sought would dictate the survival of the measure. The facts of this case were such that each variable sat at the highest extreme of each continuum: that is, the value was universally held to be of the utmost importance, the measure was fully effective and the trade-restrictiveness was absolute. This raises the question of how consistent and acceptable the new test would be when applied to a situation where the importance of the value is not as commonly accepted and the remaining variables are not such absolutes. The failure of the measure in *Korea-Beef* provides a useful comparison as the measure's efficacy was reasonable, it was not totally trade-restrictive and the value (though not stated explicitly by the AB) can safely be argued to be of relatively less importance than the protection of human health. When presented with this set of factors, the AB found that it could not justify the trade-restrictiveness of the measure.

The difficulties presented by this 'value-judging' aspect of the *Korea-Beef* test for the preservation of domestic regulatory autonomy are more clearly illustrated in the *US-Gambling* decision.[133] In this case, the US's defence was based on the public morals exception in GATS article XIV(a), which bears substantial resemblance to GATT article XX(a).[134] As such, the *US-Gambling*

[130] The *US-Gambling* shows a remarkably similar treatment of the three elements: para 6.494.
[131] ibid. [132] Neumann and Turk (n 45) 213.
[133] The GATS contains three different provisions incorporating a necessity test of differing construction: See GATS arts XIV and XIV *bis*. ArtVI:4 imposes an indirect necessity test on Members. However, only XIV has been subject to judicial attention thus far.
[134] *US-Gambling*, WT/DS285/R, paras 3.211, 6.511. The only differences are: first, that GATS art XIV(a) refers to not only the GATT XX(a) exception 'to protect public morals' but adds '*or* to maintain public order'; and second, the provision 'necessary to secure compliance with law or regulations' includes a slightly different list of goals: art XIV(c). As such, it is safe to say that the differences regarding the necessity provisions are minimal.

Panel believed that the transfer of the test to the GATS would be appropriate after also noting the common purposes between the GATS and GATT provisions.[135] The decision thus also has implications for the future application of the article XX(a) exception.

The Panel proceeded to determine whether the measure was designed to protect public morals. In doing so, the Panel took three important steps in its reasoning: first it stated that the concepts of public morals and public order were highly flexible and varied 'in time and space, depending on a range of factors including prevailing social, cultural, ethical and religious values'.[136] Secondly, it acknowledged each State's right to choose the level of protection they considered appropriate for that value.[137] Finally, however, it refrained from completely deferring the content of 'public morals' to Members, reasoning that while 'Members should be given *some* scope to define and apply for themselves the concepts of "public morals" and "public order" in their respective territories, according to their own systems and scales of values',[138] the Panel still retained the responsibility of giving meaning to these terms in order to apply them to the facts.[139] Interestingly and perhaps illogically, rather than relying on US domestic practices, the Panel proceeded to refer to the practices of other jurisdictions linking gambling issues with morality as evidence that gambling was indeed an issue of public morality.[140] The Panel's reasoning contrasts sharply with the Panel's finding in *Dominican-Cigarettes*[141] where it stated it 'finds no reason' to question the Dominican Republic's assertion that the collection of tax revenue and prevention of tax evasion is 'a most important interest for any country and particularly for a developing country'.[142] No explanation was given as to how the Panel arrived at this evaluation, illustrating a disconcerting lack of transparency in its decisions.

The *US-Gambling* Panel's reasoning was upheld by the AB, indicating that Members cannot unilaterally define public morals, potentially leading to the exclusion of certain measures that do not find international resonance despite their importance in the culture of the responding State.[143] The AB stated that the standard for the determination of necessity is an objective one. On the one hand, it stated that panels did have recourse to:

> A Member's characterization of a measure's objectives and of the effectiveness of its regulatory approach—as evidenced, for example, by texts of statutes,

[135] ibid para 6.448.
[136] ibid para 6.461.
[137] ibid. This concept is reminiscent of SPS art 5.6: see (n 149) and surrounding text below.
[138] ibid (emphasis added).
[139] ibid.
[140] ibid paras 6.471–4.
[141] *Dominican Republic-Measures Affecting the Importation and Internal Sale of Cigarettes*, WT/DS302/R, 26 November 2004 (*Dominican-Cigarettes*). The Dominican Republic claimed its tax stamp requirement for cigarette packets was 'necessary' to ensure compliance with tax and anti-cigarette smuggling laws.
[142] ibid para 7.215.
[143] Marwell (n 16) 817–19.

> legislative history, and pronouncements of government agencies or officials . . . [144]

However, it went on to state that these characterizations were not binding on the Panel without elaborating on the extent to which Members' assessments would be relevant or the standard of review that would be applied.[145] The significance of this issue was further highlighted when the AB asserted that the results of a comparison between the challenged measure and possible alternatives should be considered in light of the importance of the interests at issue.[146] The statement reinforces not only the extent to which the DBS's interpretation of the evidence on the value of the goal will dictate its survival when weighed against the other elements, making the balancing test look far more like a proportionality test *strictu sensu*, but also highlights the extent of adjudicating body discretion in determining whether the measure will even be considered to fall under one of the exceptions. The dangers posed to the preservation of regulatory autonomy as articulated in the GATT treaty are even greater in light of the continuing uncertainty regarding how each variable relates to the other and the fact that the actual balancing process is difficult to discern from each adjudicating body's reasoning.[147]

2. *The influence of the SPS and TBT agreements*

It is critical at this point to refer to the influence of the SPS and TBT necessity tests on development of the GATT balancing test. Despite the fact that the language of the tests under both these agreements provides for a LRM analysis, it appears that the adjudicating bodies have used them as guidance for introducing and clarifying aspects of the GATT balancing test largely in favour of domestic autonomy.

a) SPS agreement

It is useful to recall that the presumption under the SPS agreement is that a measure is necessary unless and until proven otherwise through the application of the LRM test contained in article 5.6.[148] This article, when read with its accompanying footnote, specifies that any proposed alternative measure

[144] *US-Gambling*, WT/DS285/AB/R, 7 April 2005, para 304.
[145] ibid.
[146] ibid para 307.
[147] Marwell notes that the balancing process by the Panel was opaque: (n 16) 813.
[148] Though art 5.6 does not explicitly refer to the word 'necessary', it has been accepted to constitute a 'necessity test': see, for example, its general treatment in *Australia-Measures Affecting Importation of Salmon*, WT/DS18/R, 12 June 1998, (*Australia-Salmon*) and Marceau and Trachtman (n 22), (n 29). Arts 2.3 and 5.5 are also highly relevant to the interpretation of art 5.6. Art 2.2 has not received any judicial consideration to date that focuses simply on whether the measure was 'necessary' but only looks at whether the measure was based on scientific evidence: *Australia-Salmon*, para 8.99.

must: (1) 'achieve the appropriate level of. . .protection', (2) be '*significantly* less restrictive to trade' and (3) be technically and economically feasible'.[149] The AB in *Australia-Salmon* affirmed that the three elements are 'cumulative' and if any one of the three is not met, the contested measure would be deemed inconsistent with article 5.6.[150] The SPS necessity test differs from its GATT and TBT cousins by explicitly removing the level of protection chosen from the ambit of judicial review. This is done by referring to the Member's 'appropriate level of protection' defined in Annex A of the agreement to mean 'the level of protection deemed appropriate by the Member'.[151] Coupled with the establishment of a higher threshold for what will be considered reasonably available (through the use of footnote 3 and the requirement that the alternative be 'significantly' less restrictive), the SPS necessity test at least on its face appears to increase the likelihood of the contested measure's survival and vests the balance of power in the hands of the responding State.

However, the drafting of the SPS necessity agreement does not firmly delineate the lines of judicial review and domestic autonomy. There is still scope for an interventionist panel to retract the apparent deferential line through a number of avenues including its determination of what degree of restrictiveness is 'significantly less restrictive',[152] the manner in which they assess whether an alternative measure will achieve a country's chosen level of protection,[153] and how the country's chosen 'appropriate level of protection' is to be discerned.[154] The limited jurisprudence on these issues has been echoed in subsequent GATT jurisprudence on necessity.

Regarding the determination of the State's chosen level of protection, the AB ruling in *Australia-Salmon* appears to tip the vertical power struggle in favour of Members.[155] It rejected the Panel's assertion that it could substitute its own reasoning about the implied level of protection,[156] finding instead that the adjudicating body had to accept the Member's expressed level of protection without regard to the level actually achieved by its chosen measure.[157]

[149] *SPS Agreement*, art 5.6, fn 3 (emphasis added).

[150] *Australia-Salmon*, WT/DS18/AB/R, 20 October 1998, para 199. This case involved Australia's import ban on uncooked salmon intended to keep fish diseases out of Australia.

[151] SPS Agreement Annex A, para 5. See also SPS arts 5.3 and 5.4.

[152] No consideration has been given to the meaning of this element. While *Japan-Measures Affecting the Importation of Apples* WT/DS245/R, 15 July 2003 (*Japan-Apples*) and *Japan-Measures Affecting Agricultural Products*, WT/DS76/R, 22 October 1998, examined and applied art 2.2, they did not rule on the issue of necessity but rather whether the measure was based on scientific evidence (though note that in *Japan-Apples* the Panel stated that the absence of scientific evidence made the measure 'disproportionate" to the risk: para 8.179).

[153] Button (n 2) 71.

[154] Case law has shown 'zero risk' to be an appropriate level if Members so wish: J Pauwelyn, 'The WTO Agreement on SPS Measure as Applied in the First Three SPS Disputes' (1999) JIEL 641, 646.

[155] Interestingly, this was held to be so without any consideration for the impact of art 5.4 requesting (not requiring) Members to take into account the objective of minimizing negative trade effects.

[156] *Australia-Salmon*, WT/DS18/R, para 8.172.

[157] *Australia-Salmon* WT/DS18/AB/R, para 199.

The only caveat was that the level had to be formulated with sufficient precision.[158] If not, the Panel could infer the appropriate level from that actually achieved by the measure.[159] This reasoning appeared to take hold in GATT jurisprudence following *EC-Asbestos*, as seen in the above analysis, and contrasts sharply with the inferred level of protection in *Korea-Beef* and *Thai-Cigarettes.*[160] It should be recalled that article XX is silent on this issue.

Further, the SPS jurisprudence has clarified that a *complaining* party must suggest a specific alternative if it is to establish a prima facie case of inconsistency with article 5.6.[161] This limitation can be contrasted with earlier applications of the GATT necessity analysis such as the *US-Gambling* Panel[162] where the Panel took an active role in suggesting alternatives not proposed by the complaining Member, raising issues about the Panel 'making the case' for that Member.[163] The extent to which this could threaten domestic autonomy was seen particularly clearly in the latter case where the Panel found that the US had not 'explored and exhausted reasonably available WTO-consistent alternatives' to the prohibition because it had rejected 'Antigua's invitation to engage in bilateral or multilateral consultations and/or negotiations . . . which could have been used to explore the possibility of finding . . .. an alternative'.[164] The extent to which this alternative was consistent with a Member's right to choose its own level of protection is highly questionable as the process of negotiation logically includes a certain expectation of compromise. Further, the mere possibility of hypothetically available alternatives that may or may not have been suggested during negotiations was deemed sufficient to defeat the measure. Thus, the Panel's reasoning appears to show greater elements of a cost-benefit analysis as seen earlier in the *Thai-Cigarettes* case rather than the balancing test as refined in *EC-Asbestos* where the AB asserted that the weighing and balancing test would not involve balancing the level of protection against the trade restriction.[165] However, the AB's rejection of the Panel's reasoning[166] coupled with its more recent treatment in *Brazil-Tyres*[167] suggests that the SPS jurisprudential approach is now also being applied under the GATT.[168]

[158] ibid para 207. [159] ibid.

[160] Note that the Panel never engaged in ascertaining the precise level of protection sought by Thailand. It simply examined alternatives believed to reduce the quantity and ensure the quality of cigarettes generally rather than to any particular extent.

[161] *Japan-Varietals*, WT/DS/76/AB/R paras 126 and 130.

[162] *US-Gambling*, WT/DS285/R, 6.531.

[163] *Japan-Varietals*, WT/DS/76/AB/R, 129.

[164] *US-Gambling*, WT/DS285/R, 6.531.

[165] See (n 123) and surrounding text above.

[166] *US-Gambling* AB expressly condemned the Panel's action in this regard: WT/DS285/AB/R, para 320. [167] *Brazil-Tyres*, WT/DS332/AB/R para 156.

[168] cf *Canada-Wheat Exports and Grain Imports*, WT/DS276/R, 6 April 2004, para 6.308 where the Panel suggested its own alternative (though note that this case occurred before both *US-Gambling* and *Brazil-Tyres*).

A final area in which the SPS necessity test has shaped GATT necessity jurisprudence is seen in the attention paid to the footnote to article 5.6 which requires alternatives to be considered with regard to their 'technical and economic feasibility'.[169] While concern has been expressed that these provisions 'invite the Panel to engage in issues of domestic resource allocation and the structure of a Member's regulatory administration',[170] its limited application indicates that it has not been used for such intrusion. Instead, it has been used to assess the true viability of proposed alternatives in the particular Member's circumstances. The special attention paid to this standard by the *Brazil-Tyres* AB is the most recent example of the adjudicating bodies' recognition of the regulatory and situational heterogeneity of WTO members,[171] and again, sits in contrast with the adjudicating bodies' decisions in *Thai-Cigarettes* and *Korea-Beef*. Thus, it is clear the SPS necessity provisions helped shape the interpretation and application of necessity under the GATT in favour of deferring to domestic autonomy. However, the effects were not immediate and are, as yet, by no means consistent.[172] Further, though some similarities can be discerned from the two tests, the compositional elements of each differ considerably.[173]

b) TBT agreement

A far closer match to the *Korea-Beef* balancing test is the TBT necessity test. Under article 2.2, the TBT prohibits States from imposing technical regulations that create 'unnecessary' obstacles to trade. Specifically, three conditions are laid down for the legitimate introduction of such regulations. First, the regulation has to pursue a 'legitimate goal': though examples are given in the provision including the protection of human and animal health, the list is non-exhaustive. Secondly, the regulation must assist in the achievement of the policy goal. Finally, the regulation must be 'not more trade-restrictive than necessary . . . taking account of the risks non-fulfilment would create'.[174] Necessity in this provision therefore seems to require some sort of balancing process but whether it imposes a proportionality test *strictu sensu* remains subject to controversy.[175] It certainly appears to provide wide scope for

[169] SPS art 5.6, footnote 3.

[170] Button (n 2) 71.

[171] See, in particular, *Brazil-Tyres*, WT/DS332/AB/R, para 156 referred to above at fn 129. See also *US-Gambling*, WT/DS285/AB/R, para 308 under the GATS. Though mentioned by the *EC-Asbestos* AB, the Panel's findings on the matter were not appealed so the AB declined to address the issue: EC-Asbestos, WT/DS135/AB/R, 169.

[172] As seen by the fact that *Korea-Beef* and *EC-Asbestos* did not address the issue.

[173] cf F Garcia, 'The Salmon Case: Evolution of Balancing Mechanisms for Non-Trade Values in WTO', in Bermann and Mavroides (n 2) 150–151 who notes substantial similarity between the test and notes the direct influence of *Australia-Salmon* in the development of the GATT necessity test.

[174] TBT, art 2.2.

[175] Neumann and Turk (n 45) 218. cf, Hilf, who argues that this does amount to proportionality: (n 48) 120. Marceau and Trachtman note that the original draft at the end of the Uruguay Round included a footnote after the additional sentence of art 2.2 which read 'this provision is

judicial intervention, not only by awarding the adjudicating body the responsibility of deciding whether the impediment to trade outweighs the consequences of non-fulfilment but, more problematically, providing scope for it to also determine whether policy objectives not falling within the explicit examples are legitimate. Thus, the tripartite criteria bear remarkable similarity to the three-prong balancing test.

The resemblance could be logically explained by the fact that, like TBT article 2.2, GATT article XX(d), to which the balancing test was first applied, both have similarly open-ended lists of legitimate policy goals and thus may make the transfer of the weighing and balancing test appropriate to some extent.[176] This is because, like XX(d), a wide variety of laws and regulations are caught by the provision thus sometimes requiring an investigation into 'the relative importance of the common interests or values that the law or regulations . . . is intended to protect'.[177] Thus while it is useful to consider the TBT jurisprudence on the issue, the fact that it has not been subject to direct judicial analysis[178] limits this exercise. The *EC-Sardines*[179] case on article 2.4 provides the only suggestion of how it will be treated. The case involved EC's prohibition on the use of the term 'Peruvian sardines' on tins containing sardine-like fish caught off the Peruvian coast which Peru claimed to be inconsistent with TBT Article 2.4.

The issue of necessity was considered following the Panel's decision (upheld by the AB)[180] that 'legitimate objectives' referred to under Article 2.4 must be interpreted in light of article 2.2.[181] The Panel's reasoning distinguished between adjudication of the policy choice as a legitimate objective and the means chosen to achieve it. It held that the wording of Article 2.2 coupled with the TBT preamble 'accords *a degree* of deference with respect to the domestic policy objectives which Members wish to pursue'.[182] It also found that the provisions show 'less deference' to the means the Member employs to achieve the objective.[183] However, in the next paragraph, the Panel went on to state that in considering the 'legitimate objectives pursued' under article 2.4, '[p]anels are . . . required to determine the legitimacy of

intended to ensure proportionality between regulations and the risks non-fulfilment of objectives would create'. Though the footnote was removed, the Note from the Secretariat affirmed that the degree of restrictiveness in the context of standard setting 'should be proportional to the risk of non-fulfilment . . . in the case of the TBT': (n 22) citing Document TER/W/16 and corr 1, (n 83) and (n 84).

[176] Neaumann and Turk (n 45) 219.

[177] ibid, 220 citing *Korea-Beef*, WT/DS161.169/AB/R,162.

[178] Though raised in a number of cases such as *EC-Asbestos*, *EC-Measures Affecting the Approval and Marketing of Biotech Products*, WT/DS291.292.293/R, 29 September 2006 and *EC-Protection of Trade-Marks and Geographical Indications for Agricultural Products and Foodstuffs* WT/DS290/R. 15 March 2005, the TBT necessity provisions were deemed inapplicable or unnecessary to consider.

[179] *EC-Trade Description of Sardines*, WT/DS231/R, 29 May 2002, WT/DS231/AB/R, 26 September 2002 (*EC-Sardines*).

[180] ibid, WT/DS231/AB/R, 286.

[181] *EC-Sardines*, WT/DS231/R, para 7.118.

[182] ibid para 7.120, (emphasis added).

[183] ibid.

those objectives',[184] which was upheld by the AB.[185] There is little, if anything, in the text of article 2.2 to prevent an adjudicating body from applying the article 2.4 logic to it.

On one reading of the Panel's reasoning, it appears that they retracted the Panel's earlier distinction between the measure and the means employed.[186] On another, closer attention to the use of the term 'degree' helps reconcile the two paragraphs. That is to say, while the Panel believed it necessary to adopt a deferential approach to considering the legitimacy of an objective, it did not believe that the choice of policy objectives was an absolute right held by states.[187] It appears to have arrived at this conclusion by referring to the fact that the TBT preamble limits Member autonomy by requiring that such measures do not constitute unnecessary obstacles to trade.[188] However, by the same token, it appears that the Panel restricted its own interference level by explicitly restricting its frame of reference to be whether the objective is 'supported by relevant public policies or other social norms'.[189] This reasoning was echoed again in the consideration of whether gambling restriction measures were caught by the public morals exception under the GATS in *US-Gambling* and highlighted the adjudicating bodies' concern that the open-ended list of objectives exposes the article to abuse by governments. It appears then that, though the TBT's influence on the directional development of the *Korea-Beef* balancing test is limited by the current lack of jurisprudence on the issue, the impact thus far has been to tip the balance in favour of trade-liberalization by awarding the adjudicating bodies more avenues of review than expressly provided for in the GATT.

However, the extent to which the TBT and SPS jurisprudence is transferrable to the GATT is limited. This is because the texts themselves differ so greatly: while TBT article 2.2 and SPS article 5.6 specifically require that technical regulations 'not be more restrictive than necessary' (though the SPS replaces 'necessary' with 'required') the GATT incorporates no such reference. The difference strongly suggests that the interpretation of the word 'necessary' under the WTO is not synonymous with the least-restrictive means test, because if the opposite were true there would have been no need for the Members to draft the TBT or SPS provisions any differently from the GATT article XX.

Despite the different textual definitions of necessity in the various agreements, the jurisprudence appears to be evolving in such a way that developments under one agreement are used to influence the development of the test under another, not merely through analogous reasoning but through the

[184] ibid para 7.121.
[185] *EC-Sardines*, WT/DS231/AB/R, para 286.
[186] See Neumann and Turk, above (n 45) 218.
[187] Camilleri, above (n 97) 26.
[188] *EC-Sardines*, WT/DS231/R, para 7.120.
[189] ibid para 7.121 referring to *Canada-Patent Protection of Pharmaceutical Products*, WT/DS114/R, adopted 7 April 2000, para 7.69.

complete transposition of criteria. While some commentators welcome the development of a more harmonized test in the interests of legal certainty,[190] it may be that the adjudicating bodies have taken the advantages of consistency too far towards uniformity, diminishing if not nullifying the different balances between regulatory autonomy and trade liberalization negotiated under each agreement, particularly the GATT.

IV. JURISPRUDENCE V THE TREATY TEXT: CONSISTENCY OR DIVERGENCE?

A. Domestic Regulatory Autonomy, Article XX, and the Korea-Beef test

The evolution of the necessity test thus far appears to be characterized by a continuing uncertainty regarding the proper delineation of power between Members and the adjudicating bodies with immense consequences for the trade/non-trade debate. While many of the adjudicating bodies' decisions should be commended for their deferential approach in appropriate circumstances, the continuing failure to develop a coherent legal doctrine on the degree of deference to be afforded to Members renders policy measures vulnerable to the adjudicating body's discretion to expand its jurisdiction through its choice of interpretative tool when applying the necessity test. The question to address now is to what extent the current balance is supported by the treaty text.

Given that the *Korea-Beef* balancing test has been recognized to still constitute an LRM test, though supplemented by a balancing analysis, the critique regarding the LRM's lack of consistency with the treaty text outlined above[191] is equally applicable. However, the 'balancing' aspect of the test still requires examination. The first and last limbs of the *Korea-Beef* test will be assessed first as it is my contention that these elements constitute the point of greatest departure from the express content of the treaty text.

1. The importance of the value and its impact on trade

To recall, the AB introduced the balancing test to the existing LRM test in its consideration of article XX(d). The wording used by the AB was very specific, stating that:

> [A] treaty interpreter assessing a measure *claimed to be necessary to secure compliance of a WTO-consistent law or regulation* may, *in appropriate cases*, take into account the relative importance of the common interests that the law or regulation to be enforced is intended to protect.[192]

[190] Art 3.2 of the DSU specifically endows the dispute settlement system with the responsibility for providing 'security and predictability to the multilateral trading system'. Note, however, that the Working Party on Domestic Regulation recently reiterated that the necessity tests cannot be used interchangeably: *WTO Secretariat WPDR Note* '"Necessity Tests" in the WTO', S/WPDR/W/27, 2 December 2003.

[191] Part IIA(2).

[192] *Korea-Beef*, WT/DS161.169/AB/R, para 162 (emphasis added).

Thus, the AB's words are highly qualified. They envisage the assessment of the value being pursued only in relation to measures 'claimed to be necessary under article XX(d)'. The difficulties presented by an open-ended list of policy objectives to the adjudicating bodies were touched upon by the Panel in relation to TBT article 2.2.[193] Under both article XX(d) and TBT article 2.2, it appeared that concern to reconcile the competing interests embodied in the agreements required some form of review to minimize its potential misuse for disguised or unjustifiable trade restrictions. However, while there appears to be some weight to this logic when applied to article XX(d), (the extent to which will be explored below), it is clear that there is no logical or textual basis for allowing the transfer of the *Korea-Beef* necessity test to article XX (b) or (a). The specific enunciation of a closed list of legitimate objectives shows that the weighing and balancing process took place at the negotiation stage of the GATT's creation. Each exception represents a policy goal that the WTO community has accepted should remain uncompromised by the pursuit of trade liberalization. Thus, provided the measure falls within one of these legitimate exceptions, the goal's relative value is beyond review.

The current test's failure to recognize the balance struck by the Members in this important regard has led to its description as 'an anathema to judicial restraint and national sovereignty'.[194] The AB's self-declared entitlement to assess the relative importance of the interests or values being pursued affords them the discretion to strike down policy measures they consider provide disproportionate non-trade benefits at the expense of free trade. The test appears to sit at odds with the purported right for Members to choose their own level of protection.

Less evident is the scope of review that can or should be given to adjudicating bodies dealing with open-listed exceptions. To give Members absolute deference over this matter would appear to shield protectionist measures from adjudicative scrutiny. However, two important issues arising from the language of the text suggest that not only is this balance of power precisely intended by the drafters but also the risk that illegitimate protectionist measures would succeed is minimal. First, and specific to the language of article XX(d), the text allows measures necessary to secure compliance with only GATT-consistent laws or regulations. I suggest that this specification should be interpreted to be just as exhaustive as the exceptions contained under sub-clauses (a) and (b). Thus, if the responding Member can show that these laws or regulations are not inconsistent with the GATT,[195] the measure will succeed. If not, the measure will fail. At no point is the Panel mandated to enquire into the relative value of the goal being pursued. This reading would not only avoid concerns regarding the adjudicating bodies' capacity and

[193] See (n 181) and surrounding text.
[194] Marceau and Trachtman (n 22) 850.
[195] *Dominican-Cigarettes*, WT/DS302/R, para 7.209.

legitimacy to second-guess the policy choices of sovereign Members, but also achieve a greater degree of transparency and consistency in the adjudicating process. I do not suggest that this measure would not be subject to significant scrutiny; instead, it is more appropriate for such examination to occur when applying the article XX chapeau. The AB in *Korea-Beef* appears to have been supporting this interpretation when it distinguished between the process of assessing whether the measure was 'necessary to secure such compliance' and whether the measure was '*designed* to secure compliance with laws and regulations that are not themselves inconsistent . . .'[196]

The content of article XX's chapeau provides strong textual evidence that no balancing test was envisaged in the interpretation of necessity. This is because the chapeau fulfils this specific function by requiring that the measure does not constitute a disguised, arbitrary or unjustifiable restriction on trade. As outlined in the LRM analysis, to allow a cost-benefit, proportionality or balancing analysis at the stage of enquiring into the necessity of a provision would render the analysis under the chapeau superfluous.[197] However, the balancing test provided for in the chapeau does not consist of the same elements as the *Korea-Beef* balancing test. Instead, the chapeau:

> . . . embodies the recognition on the part of WTO Members of the need to maintain a balance of rights and obligations between the right of a Member to invoke one or another of the exceptions of Article XX, specified in paragraphs (a) to (j), on the one hand, and the substantive rights of the other Members under the GATT 1994, on the other hand.[198]

Importantly, the AB in *US-Shrimp* went on to state that the interpretation of the chapeau marks out a line of equilibrium so that the competing rights do not 'cancel out the other and thereby distort and nullify or impair the balance of rights and obligations constructed by the Members themselves in that agreement.'[199]

The AB's words are highly significant for the current debate. By viewing the chapeau as the mechanism by which each right is prevented from 'cancelling out the other', they implicitly affirm that the rights of Members to invoke these exceptions are of equal value to the pursuit of trade liberalization. Further, they confirm that the benchmark for striking an acceptable balance under the agreement is whether it accords with the one reached by the Members themselves. The conclusions that follow from the AB's own statements are twofold. First, the value of the goal is a matter of political compromise reached between the Members when drafting each agreement, thus removing such considerations from the scope of international adjudication. Secondly, as the balancing test would distort the balance reached by

[196] *Korea-Beef*, WT/DS161.169/AB/R, para 157 (emphasis added).
[197] See (n 34) and following.
[198] *US-Shrimp*, WT/DS58/AB/R, para 156.
[199] ibid para 159.

the members by allowing the adjudicating bodies to compromise or inflate the scope of autonomy in favour or at the expense of trade liberalization at its own discretion, it is not consistent with the treaty. The balancing test implies a hierarchy of values (with trade liberalization at the top) simply not provided for in the treaty.[200]

The above analysis is equally applicable to the *Korea-Beef* balancing test's second requirement, the measure's trade-restrictiveness. This is because it is inextricably related to an assessment of the relative value of the goal. In furtherance of this analysis, and as mentioned in the LRM critique above,[201] any assessment of the extent of the goal's trade-restrictiveness is not supported by the text as this would ignore the character of the necessity provisions as exceptions to the GATT obligations.

2. The contribution of the measure to the goal

The only element of the *Korea-Beef* test that remains is the requirement that the measure contributes to the stated goal. I suggest that this is the only textually supported element of the balancing test. Returning to the text of the treaty, the ten exceptions under article XX each denote a 'different degree or connection' between the stated goal and the contested measure, as recognized by the AB itself.[202] When viewing the 'necessary' provisions in context of the other exceptions, it is clear that the term is used only to require the level of connection between the policy goal and the achievement of the policy goal be somewhat closer than the level contained in the other exceptions. It would be odd if the opposite was true: to imply a balancing test into the meaning of necessity and not any other of the end-means measurements would require not only a closer connection between the measure and goal (which is supported by the treaty text) but also permit the adjudicating body to assess the relative value of human health but not, for example, the relative value of the products of prison labour. As all the policy goals under article XX are specifically pronounced as legitimate exceptions to the general GATT obligations, this interpretation of necessity is textually untenable.

The distinction between the design of the measure and the manner of its application as recognised in *US-Shrimp* adds further weight to the argument that the consideration of necessity is simply an exercise in assessing the relationship between the ends sought and the means chosen. By reaffirming that

[200] ibid. Though the AB also mentioned that the balance was not fixed, in order to maintain consistency with its previous statements, this can only be interpreted to refer to the need for adjudicating bodies to be flexible in their application of art XX rather than providing the Panel or AB with the jurisdiction to shift the balance as it sees fit. Given the circumstances of the case (where the facts challenged the adjudicating bodies to assess the meaning of 'exhaustible natural resources' under art XX(g)), flexibility was required in order to give effect to the balance contained in the text of the treaty.

[201] Part IIA(2).

[202] *US-Gasoline*, WT/DS2/AB/R, 18.

the manner of the measure's application is only assessed when considering the chapeau, and the design at the stage of determining whether the measure falls within one of the article XX exceptions, it follows that the interpretation of necessity was simply intended to be an inquiry into the purpose of the measure and the extent to which it assists the attainment of that goal.

The texts of the SPS and TBT agreements confirm this interpretation of necessity. Of particular assistance is SPS article 2. Article 2.2 resembles the sub-clauses of article XX by requiring that the SPS measure is applied only to the extent necessary, while article 2.3 echoes the words of the article XX chapeau. The distinct separation of these requirements illustrate that whether something is necessary or not is determined purely by reference to the relationship between ends and means. Further, the explicit reference in both SPS article 5.6 and TBT article 2.2 to the least-restrictive means test demonstrates that necessity should not be interpreted as incorporating any interpretative tool unless explicitly provided for in the text. If the opposite were true, neither agreement would have needed to specify that the LRM test was required as it would have simply been inferred. Instead, the drafters chose to give effect to the marginally different balance between free trade and regulatory autonomy under the GATT, SPS and TBT by specifically incorporating the LRM test. However, it should be noted that the reference to the LRM test under the SPS and TBT is not a loose LRM test but rather an LRM test as originally understood. This is because, like the exceptions under the GATT, the value of the objectives has been determined by political negotiation and thus is beyond review.

B. The End of Balancing?

Three of the more recent considerations of the necessity test, namely in *Brazil-Tyres* and *Dominican-Cigarettes*, hint at a fundamental shift in the content of the necessity test to only constitute an examination of the contribution of the measure,[203] though subsequent cases show that any preference for change has not yet been embraced.[204] Worthy of note is the manner in which the *Dominican-Cigarettes* Panel expressed its task under article XX(d). After addressing the measure's impact on trade and the goal's relevant value, it stated, 'having said that, the Panel will focus its analysis on whether (the measure) . . . *is in fact necessary* to secure compliance with . . . (the goal)'.[205]

This statement suggests that the first two elements could be said to lie outside the question of whether the measure is necessary. While this distinction was not explicitly followed by either of the adjudicating bodies in

[203] See also, *Colombia-Ports of Entry*, WT/DS366/R.
[204] See, for example, *China-Publications and Audiovisual Products*, WT/DS363/R.
[205] *Dominican-Cigarettes*, WT/DS302/R, 7.215, (emphasis added).

Brazil-Tyres, the vast majority of the judicial reasoning in both cases was dedicated to the contribution of the goal to the intended measure. Consideration of the value of the measure and its trade-restrictiveness was limited in the Panel decision and given only a sentence each in the Appellate Body report.[206] This short string of jurisprudence provides a useful example of how a necessity test paying greater respect to the language of article XX would operate.

However, before turning to examine the necessity test in *Brazil-Tyres*, it is important to recall that nothing in the language of article XX suggests that Members did not envisage awarding any power to the adjudicating bodies in applying the necessity provisions. On the contrary, in order to maintain the equilibrium between competing goals, the assessment of necessity appears to be a matter for judicial review. In this regard, cross-reference to GATT article XXI is useful. Containing the general security exceptions, this article is littered with references to the Member's capacity to decide whether 'it considers [a measure] to be necessary'.[207] Article XXI does not impose any additional threshold. In contrast, article XX's chapeau combined with the absence of any self-judging references strongly suggests that question of necessity was intended to be reviewed by third parties.

The difficulties encountered by the adjudicating bodies in *Brazil-Tyres* when examining the level of contribution or nexus required between the ends pursued and means adopted, show that there remains a large and important role for the DBS to play in the future interpretation and application of necessity under the proposed text-based interpretation. This case concerned Brazil's measures relating to the prohibition on the importation of retreaded and used tyres which Brazil claimed was necessary under article XX(b) for reducing 'exposure to the risks to human, animal and plant health arising from the accumulation of waste tyres'.[208] Further, Brazil claimed that an exception to the prohibition which it afforded to Mercosur countries was necessary under article XX(d).

The main point of contention on appeal was the Panel's decision to apply a 'qualitative' analysis of contribution as opposed to a 'quantitative' one. This was contested by the EC who argued that the 'very indirect' relationship between the importation of retreaded tyres and the health risks at issue required a more 'diligent' examination.[209] The AB upheld the Panel's choice of methodology arguing that the logic of allowing the *risk* to human health to be assessed either quantitatively or qualitatively also applied to the analysis of the *contribution* made by a measure to the reduction of that risk.[210]

[206] *Brazil-Tyres*, WT/DS332/R, 7.108-7.114; WT/DS332/AB/R para.144; Van Calster also notes the absence of a substantive weighing process: G Van Calster, 'Faites vos jeux- Regulatory Autonomy and the World Trade Organisation after *Brazil-Tyres*' (2008) 20 J Environ L 1 121, 133.

[207] See generally GATT art XXI.

[208] *Brazil-Tyres*, WT/DS332/AB/R, para. 144, citing the Panel Report at para 7.102.

[209] *Brazil-Tyres*, WT/DS332/AB/R, para 137.

[210] ibid para 146.

The AB held that in order to be considered necessary, the measure must 'materially contribute' to the intended goal.[211] Even at first glance, this threshold suggests an inconsistency with the *Korea-Beef* AB's determination that the measure must be 'closer to the pole of indispensable than merely making a contribution' to the policy goal. This concern is further exacerbated by the AB's clarification, stating that the measure could succeed not only if evidence shows 'a measure's contribution to the goal in the past or present, but also if it *is apt to produce* a material contribution . . .'[212] The AB emphasised that just because a contribution is not immediately observable does not mean that it cannot be justified under article XX(b).[213] After considering the Brazilian strategy to achieve its goal, it considered the import ban to constitute a key element of this strategy and thus 'likely to bring a material contribution to the achievement of its objective'.[214]

On one hand, the AB' s reasoning may be argued to fall foul of consistency with the treaty text as this lower threshold diminishes the distinctions between the ten sub-clauses of article XX. The 'material contribution' test may be said to bear closer resemblance to the relationship required under sub-clauses (c) ('relating to') or (f) ('imposed for'), especially when considering how it was actually applied. It appears that the AB's earlier considerations of the ordinary meaning of necessary would assist in bolstering the legitimacy of this new threshold and its longevity. On the other hand, the reasoning demonstrated a concerted effort to take account of the complexity of regulatory environments. It recognized that a Member's choice of measure is generally informed by the broader contextual environment of culture, custom and economics and that regulatory measures do not operate in simple isolation where cause and effect is easily discerned. Instead, its operation must be observed and examined as an integral part of this broader context.[215] In many respects, the reasoning represents a climax in the development of article XX jurisprudence as it seeks to give effect to the trade/non-trade and Member/adjudicating body equilibrium constructed in the GATT. By demonstrating how an objective consideration of necessity can take account of the economic, cultural and social needs and desires of its Members, the *Brazil-Tyres* AB reconciles the need to afford sufficient deference to Member policy choices while ensuring this right is not abused. This is assisted by the fact that the value and trade-restrictiveness of the measure were never substantively assessed. The measure's subsequent failure under the chapeau in *Brazil-Tyres* shows that despite the less intrusive application of the necessity test, the two-stage analysis set out in article XX is sufficient to capture

[211] ibid para 150.
[212] ibid para 151 (emphasis added).
[213] ibid.
[214] ibid para 155.
[215] In rejecting the possible alternatives suggested by the EC, the AB noted that these measures 'already figure as elements of a comprehensive strategy designed by Brazil to deal with waste tyres. Substituting one element of this comprehensive policy for another would weaken the policy by reducing the synergies between its components, as well as its total effect': ibid para 172.

measures that have unjustifiable discriminatory effects. The decision entrenches deference as a fundamental guiding principle in not only the interpretation of the treaty provisions but also in determining the facts of the case.

It is worthy of note that the Appellate Body recently issued a landmark ruling under the SPS agreement in *US/Canada-Continued Suspension*.[216] This ruling, though limited to the interpretation of article 5.1 of the SPS agreement, enunciated the proper standard of review to be one that affords clear deference to the findings of domestic authorities with regard to 'risk assessment'. The decision suggests that we may be witnessing a broader movement towards the application of a more deferential standard of review across all the agreements. Further, given the history of cross-fertilization between the SPS and GATT agreements, as well as the fact that the SPS agreement is an elaboration of article XX(b) of the GATT, the decision may have implications for the future interpretation of the GATT necessity test. The degree of deference applied in *Brazil-Tyres* and explicitly enunciated in *US/Canada-Continued Suspension* may provide further impetus for re-form of the necessity test to one constructed in greater harmony with the text of article XX, though it remains to be seen whether this will in fact be the effect.

C. Balancing, LRM and the WTO: A Possible Future?

Finally, it is useful to briefly consider whether it is desirable to incorporate the two dominant judicial tools employed when interpreting necessity. The application of both the balancing test and the strict LRM test has received substantial support by different commentators. First, the LRM test (as defined in Part I as opposed to the loose LRM test which was suggested to have been applied in Part II) is said to offer the benefits of transparency and predict-ability by requiring a 'particularised comparison between the measure in question and a specific proposed alternative'.[217] The test is thereby said to eliminate inefficiently trade-restrictive measures and avoid the ambiguities presented by allowing the adjudicating body to question the value of the goal.[218] However, if the measure is instrumental to the achievement of the goal (that is, if it makes more than simply a contribution) it is difficult to see how an alternative measure with less impact on trade could achieve the same level of protection sought by the Member. The application of the LRM test thus far demonstrates a potential danger of the test being used in a manner that compromises or explicitly rejects the Member's chosen level of

[216] *United States-Continued Suspension of Obligations in the EC-Hormones Dispute*; WT/DS320/AB/R, *Canada-Continued Suspension of Obligations in the EC-Hormones Dispute*, WT/DS321/AB/R.
[217] Marwell (n 16) 829.
[218] ibid 828.

protection.[219] Thus, until the adjudicating bodies can show how this can be avoided, I suggest that its application under the GATT nullifies the suggested benefits by undermining the protection afforded to domestic regulatory autonomy contained in the text. The validity of this concern is supported by the presence of safeguards for the level of protection expressly contained in the SPS and TBT provisions, thus minimising the potential threat posed by the LRM test.

The *Korea-Beef* balancing test has also been seen to assist the WTO adjudicating bodies to give effect to the dual objectives, mainly by providing greater flexibility to the assessment process.[220] This flexibility is said to allow a greater number of factors to be taken into account and thus 'eases the justificatory burden for a member to protect its violating regulations'.[221] However, the degree of discretion the test affords to the adjudicating bodies is the greatest point of concern for the adjudication of non-trade values in a consistent and transparent manner. This concern has manifested itself on numerous occasions as demonstrated in the comparative analysis of the test's application.[222] Given the sensitive issues such balancing tests need to consider, a strong sense of democratic legitimacy needs to be present before a body purports to undertake any such exercise.[223] Neumann and Turk point to a number of ongoing problems facing the WTO, suggesting that it is not yet ready to begin employing balancing tests, though they specifically refer to the inappropriateness of a proportionality test.[224] Amongst others, they cite the access problems facing developing countries and other stakeholders in the formation and application of rules as well as the fact that 'not even all the major cultures and legal families seem to be adequately represented in the Appellate Body'.[225] This can be contrasted not only with the ECJ in which every State has a judge of its own nationality, but also other international courts with similarly large numbers of Members such as the ICJ and ITLOS, which both allow for introduction of ad hoc judges.[226]

The current lack of predictability or certainty in the test's application, coupled with the opaque application of the balancing test strongly suggest that the introduction of balancing processes where the value of the goal is assessed relative to the impact on trade by the adjudicating body is premature at the very least. The protracted negotiations that have taken place over whether to introduce an expanded non-exhaustive list of exceptions and/or the introduction of a proportionality test to GATS VI:4[227] demonstrate not only how difficult these issues are to resolve in accordance with the aims of the agreement but also that the decision is ultimately a product of political negotiation, not a judicial one.

[219] See *Korea-Beef, Thai-Cigarettes, US-Gambling*.
[220] Osiro (n 84) 140.
[221] ibid.
[222] See Part II generally.
[223] Neumann and Turk (n 45) 232.
[224] ibid 232. Howse and Turk (n 2) 326.
[225] ibid.
[226] ibid.
[227] See Kennet et al (n 64) 6–9.

V. CONCLUSION

The need to ensure that the domestic regulatory freedom contained in the treaty is not abused by Members is logically an issue of concern for the future viability of the WTO system. However, this need has seen the adjudicating bodies develop the necessity tests in a manner that distorts the trade liberalization/domestic autonomy equilibrium contained in the treaty. By expanding the adjudicating bodies' jurisdiction and discretion through the use of the LRM and balancing tests, significant controversy has been generated over the adjudicating bodies' competence and legitimacy to weigh the value of regulatory goals against their negative trade effects.

I have argued that as the treaty text provides adequate safeguards for the proper review of the uses of the exception clauses by Members and as the freedom to pursue certain goals has already been agreed upon by the Members during the negotiation of the treaty, the continued operation of both the LRM and *Korea-Beef* balancing tests are not only textually unsupportable but an illegitimate and unwarranted intrusion into Members' autonomy. Instead, I have proposed an alternative test that gives proper effect to the article XX sub-clauses and chapeau and offers a more transparent legal standard against which these measures can be assessed. Finally, and most significantly, the test meets the challenge of eliminating protectionist measures while preserving domestic regulatory autonomy. The task of identifying illegitimate protectionist measures disguised under the necessity provisions clearly requires independent, and often relatively intrusive, adjudication. However, this must be done in a manner that pays due respect to the negotiated compromise achieved by the Members. Such recognition is imperative if the WTO adjudicating bodies are to retain their legitimacy and the Members' cooperation.